## Environment Canada Imaging Cover Page

Report N.:



SKP Box Number:

672672428

BOX 159 TORONTO AME, ONTARIO

ATMOSPHERIC ENVIRONMENT SERVICE DEPARTMENT OF THE ENVIRONMENT – CANADA

# Technical Memoranda

AN ANALYSIS OF OBJECTIVE WINTERTIME FORECASTS OF PRECIPITATION PROBABILITY AT SIX AIRPORTS IN BRITISH COLUMBIA

by

### M. ROSE and L. PARENT

TEC 779 11 OCTOBER, 1972

U.D.C. 551.509.324.2

### ENVIRONMENT CANADA - ATMOSPHERIC ENVIRONMENT SERVICE 4905 Dufferin Street Downsview, Ontario

### AN ANALYSIS OF OBJECTIVE WINTERTIME FORECASTS OF PRECIPITATION PROBABILITY AT SIX AIRPORTS IN BRITISH COLUMBIA

by

#### M. Rose and L. Parent

#### ABSTRACT

A program has been developed to prepare objective probability of precipitation forecasts at six airports in British Columbia. An explanation of its development and results of a three month operational run are discussed.

### UNE ANALYSE DES PRÉVISIONS OBJECTIVES D'HIVER SUR LES PROBABILITÉS DE PRÉCIPITATIONS À SIX AÉROPORTS DE COLOMBIE-BRITANNIQUE

#### par

M. Rose et L. Parent

### RÉSUMÉ

Un programme a été mis au point pour préparer les probabilités objectives sur les prévisions de précipitations à six aéroports de la Colombie-Britannique. On y étude une explication de sa mise au point et les résultats de trois mois d'operation.

### AN ANALYSIS OF OBJECTIVE WINTERTIME FORECASTS OF PRECIPITATION PROBABILITY AT SIX AIRPORTS IN BRITISH COLUMBIA

by

#### M. Rose and L. Parent

#### (Manuscript received June 27, 1972)

#### 1. Introduction

Parent and Rose(1) have reported on the use of a multivariate discriminant analysis (MDA) technique to predict probability of precipitation during the summer at Vancouver and Prince George Airports. This report extends the study to the 'winter months' of February, March and April, and to four additional locations, the airports at Cranbrook, Revelstoke, Prince Rupert and Penticton.

#### 2. Predictors in Discriminant Analysis

As reported earlier, the dependent variables were centered twelve hour precipitation amounts and mean centered twelve hour cloud amounts. The probability of occurrence of cloud and precipitation was forecast for the six airports listed above. In all, 83 predictors were screened to determine those best suited to the MDA technique at these places. Prognostic data as well as actual data were considered in this process. The former included vertical velocity and relative humidity taken from the U.S. primitive equations (PE) model. Predictors of the second type included initial values of sea-level pressure and 500 mb. height, sea-level pressure gradients as measured by pressure differences at some 14 observing sites, and 500 mb. wind speed and direction at seven upper air stations 'upstream', near or in British Columbia. North-south and east-west components of these winds were also considered. Ten predictors were chosen for each station and discriminant analysis was made on the basis of 'rain' or 'no rain' and 'cloud' or 'no cloud'. the latter being defined as less than seven tenths of low and middle cloud.

#### 3. Results on Dependent Sample

The dependent data set covered the period February-April 1971. Prediction equations were derived for both parameters at Vancouver, Prince George, Penticton, Cranbrook and Prince Rupert. At Revelstoke the precipitation MDA was done in conjunction with a Rogers Pass snowfall study and cloud data were not available in time to prepare equation for use in the operational runs. At Prince George, discriminant functions were developed for cloud and these were included in the operational runs. Unfortunately, restricted use of the computer prevented a printout of the individual results on the dependent data set, and hence no contingency table could be prepared.

The results of an analysis of performance on the dependent data set in the form of contingency tables with percentage correct are in Appendix 1. MDA probabilities of 50% and over were verified as forecasts of occurrence, 49% or less as non-ôccurrence for both cloud and precipitation. Skill scores as defined by Brier and Panofsky (2) for the dependent sample are in Appendix 6.

The standard MDA significance statistic, the Mahalanobis  $D^2$ , as discussed by Miller (3) was computed for all cloud and precipitation functions and a table of these results is in Appendix 5. The generalized Mahalanobis  $D^2$  can be used as chi-square (under assumption of normality) with M (g-1) degrees of freedom to test the hypothesis that the mean values are the same in all the g groups for these M variables.

In our case, the value of chi-square for 10 variables and 2 groups at the .05 significance level is 18.3 and at the .01 significance level it is 23.2. Typical values we have calculated range from 50 to near 200 and are, therefore, highly significant.

In addition, at Vancouver Airport an analysis was made on the basis of three categories. These were defined as  $0\pm2$  tenths, 3-7 tenths, and 8-10 tenths for mean cloud amount and nil, Tr-.24, and .25 or more for precipitation. This three-way analysis was intended to provide a quantitative estimate of cloud amount and precipitation in ranges that were meaningful, considering the climate of the station. In this three group MDA, the group for which the probability was highest was verified. This probability could be as low as 34% if the remaining groups had individual probabilities of 33%.

Because the contingency tables in Appendix I display results obtained by using the prediction equations on the dependent data sample, two features should be noted. The totals of the 'actual' rows give the frequency of each category in the sample. The figures in the boxes represent a near optimum level of effectiveness for the technique, using these predictors. Subsequent study has shown that a screening technique using an equal number of cases in each group can give better results in some cases and this technique will be used to derive future functions.

#### 4. Operational Use

The prediction equations derived above are used to produce guidance in real time at the Weather Office, Vancouver. Cloud and precipitation probability forecasts are produced by computer for twelve hour intervals centered 12, 24, and 36 hours from reference time. Forecast office technicians extract the data required in the programme twice daily, either from numerical prognostics or from observations at the main synoptic times of 0000Z and 1200A. This data base is also used in other office programmes. Operations technicians enter the data on paper tape at the terminal of the timesharing computer system and run the operational programme. A sample of the print-out is given in Appendix 2.

#### 5. Verification

Approximately 115 probability forecasts are prepared each 24 hours. These have been verified for the 90 day period from February 1, 1972, until April 30, 1972. The 12, 24 and 36 hour forecasts were scored separately to display the decrease in accuracy attributable to deterioration in the prognostic data fields used in the equations. Verification of these operational forecasts is given in Appendix 3. The corresponding skill scores shown are in Appendix 7. In deriving the equations, 'no cloud' was defined as 0-6 tenths inclusive. This bias was intentional since it is important in a predominantly cloudy environment to emphasize in the forecast even restricted amounts of sunshine. However, verification was based on a convenient manuscript summary maintained in the office for other purposes, and cloud amount in this summary was given only in terms of airways symbols. This had the effect of defining 'no cloud' as 0-5 tenths in the verifying data. When the technique is verified in this manner, a number of forecasts will be scored incorrectly. Since actual cloud amounts were readily available from the daily record of observations at Vancouver Airport, verification there was done in both ways, i.e., 'no cloud' was equated to 0-6 tenths of cloud as well as to 0-5 tenths. It was then possible to estimate the magnitude of the error attributable to symbolic verification. The results are given in Appendix 4; comparison with those in Appendix 3 shows, as expected, that the number of 'no cloud' occurrences is larger. The magnitude of the increase, about 30%, casts a little doubt on the effectiveness of the symbolic sky condition as a basis for verification. In all three time periods, a total of 17 forecasts of cloudy, considered correct on the symbolic definition moved into the incorrect category on the precise definition, but to compensate, 19 forecasts of 'no cloud' previously

considered incorrect moved into the correct category. Overall, the percentage of correct forecasts and the skill scores differed very little from one system to the other.

At Vancouver, use of the proper verification basis reduced the subsequent agreement from 92% to 87% when cloudy weather was forecast and raised it when non-cloudy weather was forecast from 44 to 54%. Since the sample period was predominantly cloudy, correct forecasting of non-cloudy weather took on more importance and therefore precise verification placed the results in this category in a better light. However, with the increase in the number of days defined as non-cloudy, the extent to which non-cloudy days were correctly forecast dropped from about 84% in the approximate verification to 76% in the more precise one.

Considering all airports again, in the first twelve hour period, on the average 75% of the forecasts were correctin both categories (cloud and precipitation). This compared with the average optimum figure of 79% for cloud and 77% for precipitation derived from the dependent sample. In the second twelve hour period the average in both categories drops to about 72.5% and in the final period to 67.5%. These are averages only and, as can be seen from the tables, performance varies considerably with both category and station. At any given location there is frequently much more skill in forecasts of one of the two mutually exclusive possibilities than in the other.

For the 'rainy' airports, (Vancouver, Prince Rupert, and Revelstoke) the percentage of correct forecasts in the 'rain' category was higher than in 'no rain'. The reverse was true at 'dry' airports (Penticton, Cranbrook and Prince George). The climatological probability of rain is, of course, involved in this result. The skill scores remove this climatological factor. In general, on the dependent sample, at the rainy airports, there is more skill in forecasts of 'no precipitation' even though a greater percentage of the forecasts of precipitation were correct. At the 'dry' airports, there was also more skill in the forecasts of precipitation. The same pattern is noticable on the verified sample.

#### 6. Sources of Error in the Operational Runs

In general, the forecasts showed lower skill than did the results on the dependent sample.

Possible reasons for deterioration in time are listed below:

 (a) Dependent data were "perfect progs". Actual run time data were from forecast fields.

- (b) Because of the amount of data necessary, grid point data was extracted from the prognostic forecast fields and linearly interpolated by computer.
- (c) The dependent sample 500 mb, winds were from Radiosonde data whereas run time 500 mb, winds were from forecast 500 mb, height field gradients.
- (d) The map scale of the forecast fields was 1 in 30 million and maps were very hard to read at times. Quality control checks indicated that errors could creep in due to illegible maps.
- (e) In cases of missing surface P.E. model prognostics, the Vancouver forecast prognostics for the nearest valid time were used. This was usually an earlier prognosis valid 6 hours earlier, and in cases of missing 500 mb. maps, the CAO baroclinic 500 mb. prognostic maps were used. In cases of missing vertical velocity and relative humidity maps, the mean vertical velocity and relative humidity of the dependent sample was used for each station.

Of these reasons for poorer performance, the last is most significant. The vertical velocity and relative humidity correlate most highly with cloud and precipitation and the loss of maps of these fields was a significant drawback. However, only some 5% of these maps were missing.

One must examine the distributions of predictors to find other sources of error. The MDA technique essentially sorts between groups in the independent variable based upon the spread between groups in the dependent variables or predictors. The Mahalanobis  $D^2$  is the standard test statistic used to determine if the independent variable groups are significantly well separated by the predictors to assess the performance of the technique on dependent data. The assumption of normality is that the distributions of the independent variable and the predictors are normal. This is, of course, not true for all the predictors, of for the cloud and precipitation distributions. For this reason, skill scores and per cent correct are also shown in the appendices.

7. Conclusions

It is apparent from the tables in the appendices that the skill scores on the operational runs are somewhat lower than those on the dependent sample.

The reasons for these lower scores are outlined in Section 6 (Sources of Error) above. The most significant reason, however, appears to be the inherent errors in prognostic fields. The operational results that have been scored are not a fair test of the MDA technique on independent data. It should be borne in mind that what have been scored are the results of a modified operational technique that cannot approach the scores on the dependent sample. Despite the lower skill scores on these real-time runs, the forecast probabilities mirror the forecast fields, and as such, provide an excellent first estimate for the operational forecasters. The forecasters are aware of the frequent uncertainties in forecast fields and use the probabilities in this light. A further point to note is that to score probability forecasts correctly, more sophisticated scoring systems should be used. We have scored as "correct" those forecasts of 50% probability and over that were coincident with an event and those forecasts of 49% and less that were coincident with a non-event. In fact, if the forecast probabilities are distributed normally, the "correct" forecasts based on this system should only average 75% correct.

The encouraging results on the operational runs, using in some cases substitute predictors, and in all cases, prognostic data, indicate that the MDA technique applied in this manner can be a very useful tool in aiding the forecaster.

APPROVED,

Whole

J.R.H. Noble, Assistant Deputy Minister, Atmospheric Environment Service.

- 7. References
  - 1. Parent, L. and M. Rose., 1972: Precipitation Probability Forecasting in B.C. Using a Multivariate Discriminant Analysis Technique, Environment Canada, Atmospheric Environment Service, Technical Memoranda series, TEC 774.

7

- 2. Panofsky, H. and G.W. Brier., 1958: Applications of statistics to Meteorology, University Park, the Pennsylvania State University, pp.224.
- 3. Miller., 1962: Statistical Prediction by Discriminant Analysis. Meteorological Monographs. Volume 4, No. 25.

e se 🖬

### DEPENDENT SAMPLE





#### THREE GROUP PROBABILITIES

CLD (TENTHS)

FCST

т

178

PCPN (Nil, 0-.24, Gt. OR = .25)FCST

		HVY	LGT	ŊIL	Т	_
A	HVY	18	6	0	24	
C T	LGT	15	43	9	67	
U A L	NIL	1	16	70	87	
	Т	34	65	<b>7</b> 9	178	

Percent Correct 74%

8-10 56 22 7 85 9 20 40 11 13 39 53 1

8-10 3-7 0-2

Α

C T

U А

L

3-7

0-2

Т

68

Percent Correct 70%

55

55

### DEPENDENT SAMPLE

PRINCE RUPERT AIRPORT







А

C Т

U

А

 $\mathbf{L}$ 

PRINCE GEORGE AIRPORT

PCPN FCST

**REVELSTOKE AIRPORT** 

PCPN FCST



Percent Correct 83%

PCPN N/P Т 63 83 PCPŃ 20 95 N/P 3.8 57 101 77 178 Т

Percent Correct 68%

- 9 -

\$ ..

### DEPENDENT SAMPLE

PENTICTON AIRPORT

Α

С

Т

U

Α

 $\mathbf{L}$ 







er	cent	Correct 82%	

PCPN N/P T						
PCPN	44	15	59			
N/P	30	. 89	119			
Т	74	104	178			

Percent Correct 68%

CRANBROOK AIRPORT

Α

С

Т

U

А

L

CLDFCST

PCPN FCST



PCPN N/P T							
PCPN	44	10	54				
N/P	37	87	124				
T.	81	9,7.	178				

Percent Correct 74%

### DISCRIMINANT ANALYSIS GUIDANCE BASED ON 030000Z DATA

	WED NITE		THURSDAY		THU NITE	
	Prob of Cld	Prob of Pcpn	Prob of Cld	Prob of Pcpn	Prob of Cld	Prob of Pcpn
Vancouver	0.09	0.08	0.18	0.24	0.62	
Prince George	0.40	0.52	0.52	0.72	0.62	0 <b>. 79</b>
Penticton	0.03	0.39	0.12	0.71	0.49	0.87
Cranbrook	0.19	0.18	0.26	0.26	0.47	0.50
Revelstoke	0.	0.22	0.	0.36	0.	0.83
Prince Rupert	0.54	0.37	0.70	0.61	0.80	0.87

ADDITIONAL VANCOUVER INFORMATION

CLOUD AMOUNT PROBABILITIES PRECIPITATION AMOUNT PROBABILITIES

	0-2 Tenths	3-7 Tenths	8-10 Tenths	s i s <b>Ni</b> ls	025 Inches	=.25 Inches
Wednesday Night	0.63	0.25	0.12	0.81	0.10	0.04
Thursday	0.42	0.34	0.24	0.65	0.23	0.12
Thursday Night	0.05	0.40	0.55	0.01	0.16	0.83

- 12 -

### VANCOUVER FEB-APR 1972 PRECIPITATION



FCST CLD NO CLD TOT CLD 66 57 123 A C NO CLD 6 33 39 T TOT 72 90 162

Percent Correct 61%

### VANCOUVER AIRPORT

### FORECAST VERIFICATION FEB-APR 1972

### THREE GROUP PROBABILITIES

### CLOUD (TENTHS)



24 Hr. FCST

١.



		0-2	3-7	8-10	TOT
•	0-2	7	5	0	, 12
A C T	3-7	9	24	12	45
U A L	8-10	11	44	47	102
	TOT	27	73	59	159
	•	To	tal Co	rrect 4	19%

36 Hr. FCST

	-	0-2	3-7	8-10	TOT
A C T	0-2	7	4	0	11
	3-7	12	25	9	46
U A	8-10	15	38	49	102
L	TOT	34	67	58	159

Total Correct 51%

### VANCOUVER AIRPORT FORECAST VERIFICATION FEB-APR 1972 THREE GROUP PROBABILITIES

- 14 -

### PRECIPITATION ·



24 Hr. FCST



		NIL	0-25	=,25	TOT.
•	NIL	47	1,6	4	67
ч С Г	0-25	12	53	6	71
U A L	=.25	2 ·	16	. 4	22
_	TOT	61	85	14	160

Total Correct 68%

Total Correct 65%

36 Hr. FCST

		NIL	0 <b>-</b> 25	=.25	TOT
А	NIL	41	21	5	.67
C T	0-25	19	40	11	70
U A L	=.25	1	17	5	23
L	TOT	61	78	21	160

Total Correct 54%

### SUMMARY OF FORECAST VERIFICATION

### at

Penticton, Prince George, Cranbrook, Revelstoke, and Prince Rupert

AIRPORT	ENT CORRECT FORECASTS					
	12 hr. pcpn	fcst cld	24 hr. pcpn	fcst cld	36 hr. pcpn	fcst cld
Penticton	72	71	64	72	65	65
Cranbrook	69	75	62	69	60	65
Prince George	69	<b>7</b> 5	73	78	59	69
Revelstoke	7.6	-	79.	-	72	-
Prince Rupert	87	81	82.	81	83	77

ʻ b

- 15 -





36 Hr. FCST

	CLD 1	CLD NO CLD TOT				
A CLD C	60	51	111			
	12	39	51			
A L TOT	72	90	162			

Total Correct 61%

•	PRECIPITATION									
<u>Stati on</u>	<u>Mahalanobis D<sup>2</sup></u>	Variables	Observations							
Vancouver	195	10	178							
Prince George	52	10	178							
Penticton	76	10	178							
Cranbrook	69	10	178							
Revelstoke	158	10	178							
Prince Rupert	140 .	10,	178							
		÷.,								

### VALUES OF MAHALANOBIS D<sup>2</sup> FOR DISCRIMINANT FUNCTIONS DEVELOPED ON DEPENDENT DATA SET

CLOUD

	Mahalanobis D <sup>2</sup>	Variables	Observations
Vancouver	97	10	178
Prince George	52	10	178
Penticton	105	10	178
Cranbrook	68	10	178
Prince Rupert	108	10	178
Revelstoke	(Not available)	<b></b>	a = =

	ססד	᠂ᡴ᠋ᠣᡗ᠇᠇᠔᠇ᢇ					
		NO		NO			
STATION	PCPN	PCPN	TOT	CLD	CLD	TOT	
Vancouver	.66	.72	.68	. 55	.62	.58	
Prince George	.30	. 44	.35	(Not a	vailable	)	
Penticton	. 39	.56	. 46	.60	.69	.64	
Cranbrook	.34	.66	. 45	. 40	<b>. 57</b>	. 47	
Revelstoke	.67	.62	.64	(Not A	vailable	;) <sup>1</sup>	
Prince Rupert	.65	.74	.69	. 59	.51	.54	

# SKILL SCORES FOR DEPENDENT SAMPLE FORECAST

·

. . . .

- 19 -

### SKILL SCORES FOR FEB. - APR. 1972

### OBJECTIVE FORECAST VERIFICATION

·		<u>12 Hr.</u> No		24 Hr.			<u>36 Hr.</u> No		
STATION	<u>Pcpn</u>	Pcpn	Tot	Pcpn	Pcpn	Tot	Pcpn	Pcpn	Tot
Vancouver	.43	. 62	.51	.47	.41	.44	. 42	. 39	.41
Penticton	<b>.</b> 36	.48	.41	. 22	.15	:18	.28	.24	. 25
Cranbrook	:28	.64	. 39	.17	.32	.22	.14	.30	.19
Prince George	.26	.52	. 35	.16	.33	.22	.11	.25	.15
Revelstoke	.39	.52	.44	.44	.62	.52	.34	.40	.37
Prince Rupert	.63	.71	.67	. 49	.60	. 54	.43	.47	. 45
						. :			

#### PRECIPITATION FORECASTS

CLOUD FORECASTS

	12 Hr.			24 Hr.			36 Hr.		
	No		:	No		No			
	Cld	Cld	Tot	Cld	Cld	Tot	Cld	C1d	Tot
Vancouver	.68	.28	. 39	.71	.18	. 29	.65	.17	. 26
Penticton	.34	.48	.40	. 39	.44	.41	.28	.31	. 29
Cranbrook	.42	.54	.48	.33	.43	.37	. 25	.33	. 29
Prince Rupert	.70	.36	.47	.69	.35	.46	. 69	.26	.37

### CLOUD FORECASTS VERIFIED

		(As cloudy 7-10 tenths)								
	·	<u>12 Hr</u>	•	·····	24 Hr.			<u>36 Hr.</u>		
•		Νo			No	•		No		
	Cld	Cld	Tot	Cld	Cld	Tot	Cld	Cld	Tot	
Vancouver	.58	,35	.44	, 52	.19	.28	.47	.17	.25	

	• •			
TEC 779	UDC.551.509.324.2	TEC 779	UDC.551.509.324.2	
11 October, 1972		11 October, 1972		
	CANADA		CANADA	
Environment - Atm 4905 Dufferin St	nospheric Environment Service reet, Downsview, Ontari <b>o.</b>	Environment - A 4905 Dufferin	Atmospheric Environment Service Street, Downsview, Ontario.	
An Analysis of Obj Precipitation P Br	ective Wintertime Forecasts of robability at Six Airports in itish Columbia by	An Analysis of C Precipitation	Objective Wintertime Forecasts of Probability at Six Airports in British Columbia by	
M. R	ose and L. Parent.	Μ.	Rose and L. Parent.	
6 pps. 12 tables, 3 re	fs.	6 pps. 12 tables, 3	refs.	
Subject Reference: 1. 2. 3.	Probability Forecasts Precipitation British Columbia	Subject Reference:	<ol> <li>Probability Forecasts</li> <li>Precipitation</li> <li>British Columbia</li> </ol>	
TEC 779	UDC:551.509.324.2	TEC 779	UDC:551.509.324.2	
11 October, 1972		11 October, 1972		
Environment - Atm 4905 Dufferin St	CANADA nospheric Environment Service treet, Downsview, Ontario.	Environment - A 4905 Dufferin	CANADA Atmospheric Environment Service n Street, Downsview, Ontario.	
An Analysis of Obj Precipitation Pr Br M. Re	ective Wintertime Forecasts of robability at Six Airports in itish Columbia by ose and L. Parent	An Analysis of O Precipitation M.	bjective Wintertime Forecasts of Probability at Six Airports in British Columbia by Rose and L. Parent	
6 pps. 12 tables, 3 re Subject Reference: 1. 2. 3.	fs. Probability Forecasts Precipitation British Columbia	6 pps. 12 tables, 3 Subject Reference:	refs. 1. Probability Forecasts 2. Precipitation 3. British Columbia.	

ABSTRACT: A program has been developed to prepare objective probability of precipitation forecasts at six airports in British Columbia. An explanation of its development and results of a three month operational run are discussed.

ABSTRACT: A program has been developed to prepare objective probability of precipitation forecasts at six airports in British Columbia. An explanation of its development and results of a three month operational run are discussed.

ABSTRACT: A program has been developed to prepare objective probability of precipitation forecasts at six airports in British Columbia. An explanation of its development and results of a three month operational run are discussed.

ABSTRACT: A program has been developed to prepare objective probability of precipitation forecasts at six airports in British Columbia. An explanation of its development and results of a three month operational run are discussed.